The private classical capacity with a symmetric side channel and its application to quantum cryptography

Graeme Smith[1,*]

[1]*Department of Computer Science, University of Bristol, Bristol, BS8 1UB, UK*

(Dated: May 25, 2007)

We study the symmetric-side-channel-assisted private capacity of a quantum channel, for which we provide a single-letter formula. This capacity is additive, convex, and, for degradable channels, equal to the unassisted private capacity. While a channel's (unassisted) capacity for for private classical communication may be strictly larger than its quantum capacity, we will show that these capacities are equal for degradable channels, thus demonstrating the equivalence of privacy and quantum coherence in this context. We use these ideas to find new bounds on the key rate of quantum key distribution protocols with one-way classical post-processing. For the Bennett-Brassard-84 (BB84) protocol, our results demonstrate that collective attacks are strictly stronger than individual attacks.

arXiv:0705.3838v1 [quant-ph] 25 May 2007

I. INTRODUCTION

One of the earliest results in quantum information theory was the realization in [1] that a noisy quantum channel can be used to establish secret correlations whose security is guaranteed by the fundamental laws of physics. Furthermore, while the full-scale implementation of quantum computation is likely to remain a distant hope for years to come, secure quantum key distribution protocols may begin to play an important role in the world of information security in the not-too-distant future.

In the simplest case of independent and identically distributed noise, which we will consider here and to which most quantum key distribution (QKD) protocols can be reduced [2, 3], the capacity of a channel for private classical communication was studied in [4]. In that work, a *multi-letter formula* for the private classical capacity was provided (and, indeed, a similar formula for a channel's capacity for quantum communication). Unfortunately, this multi-letter formula cannot be evaluated in general, and thus provides only a partial characterization of the capacity we seek.

In lieu of a closed form expression for the private classical capacity of a quantum channel, which we will call C_p, it is the primary purpose of this work to provide upper bounds. In particular, we will consider the capacity of a quantum channel for private classical communication when assisted by the family of (one-way) quantum channels that map symmetrically to their output and environment. Since any such channel has zero private capacity on its own, one would expect the resulting bound to be quite tight. This approach is very much in the spirit of [5], and in fact our expression for the symmetric side-channel assisted private capacity (ss-private capacity) shares many of the nice properties of the ss-capacity found there, namely it is single-letter, additive, and convex.

A secondary goal of this work is to explore the connection between unconditional privacy and quantum coherence, which has long been folklore in the quantum information community and provided motivation for the coding theorems proved in [4] (see also [6, 7]). While this analogy is quite useful, it is known that the correspondence is not exact. Indeed, there are quantum channels for which the capacity for private communicaton and quantum communication are quite different. In [8] it was shown that there exist quantum states from which no entanglement can be distilled via two-way classical communication but which nevertheless can be used to create secure key via one-way public classical communication. This leads to examples of channels with zero quantum capacity but nonzero private classical capacity.

Understanding the connection between coherence and privacy in a quantitative way does not seem to be possible at the moment, as there do not yet exist simple expressions for either the private classical or quantum capacities of a channel. However, we will show below that, for the class of channels known as *degradable*, it is possible to find a simple expression for the private classical capacity, C_p, and indeed for such channels C_p is exactly the quantum capacity (which, due to [9] has a closed-form expression). As well as giving the first examples of nontrivial channels for which C_p can be found explicitly, this provides a setting in which the above mentioned analogy between privacy and coherence can be made exact.

Furthermore, the ss-private-capacity of a degradable channel is exactly equal to its (unassisted) quantum capacity. We will combine this result with the convexity of the ss-private capacity to provide a new technique for upper-bounding the private capacity of a general quantum channel, extending the current best known bounds for the quantum capacity of the depolarizing to its private capacity and providing new bounds for private capacity of a channel with independent phase and amplitude noise. This last result leads to collective attacks on BB84 that outperform the optimal individual attack.

*Electronic address: gsbsmith@gmail.com

The rest of the paper is organized as follows. In Section II we study the private capacity of a degradable channel, in Section III we provide a single-letter formula for the ss-private capacity of a general channel, while in Section IV we provide upper bounds for the private capacity of some specific quantum channels and discuss their relation to collective attacks in QKD. In Section V we mention a few open problems.

II. NOISY PROCESSING IS NO HELP FOR DEGRADABLE $\mathcal{N}$

In a classical setting, if we imagine a broadcast channel which maps $\mathcal{N} : X \to (Y, Z)$, where Y is the output to the receiver and Z is the output of an eavesdropper, it was shown in [10] that the secret-key capacity of $\mathcal{N}$ is exactly

$$C_p(\mathcal{N}) = \sup_{X \to T} \left(I(T;Y) - I(T;Z) \right). \tag{1}$$

Here the optimization is over a reference variable X, which represents the distribution of messages sent through the channel, together with a noisy processing of X that generates T.

By analogy with this result, one may imagine that the private classical capacity of a quantum channel would be given by

$$C_p^{(1)}(\mathcal{N}) := \sup_{\{p_x, |\varphi_x\rangle\}, X \to T} \left(I(T;B)_\omega - I(T;E)_\omega \right),$$

where $\omega_{ABE} = \sum_{x,t} p(t|x)p(x)|t\rangle\langle t|_A \otimes U_{\mathcal{N}}|\varphi_x\rangle\langle\varphi_x|U_{\mathcal{N}}^\dagger$ with $U_{\mathcal{N}}$ an isometric extension of $\mathcal{N}$ (i.e., $\mathcal{N}(\rho) = \mathrm{Tr}_E\, U_{\mathcal{N}}\rho U_{\mathcal{N}}^\dagger$). So, the optimization would again be taken over input random variable X (this time with a choice of basis), together with a classical noisy processing $X \to T$. Indeed, the coding theorem proved in [4] showed that this rate is in fact achievable:

$$C_p(\mathcal{N}) \geq C_p^{(1)}(\mathcal{N}),$$

but did not establish the converse statement. Instead, it was shown that

$$C_p(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} C_p^{(1)}(\mathcal{N}^{\otimes n}).$$

Evidence was found in [11] that this regularization, as the limit over n is typically called in this context, is necessary in general.

A class of channels for which we will be able to explicitly evaluate C_p are called degradable, and were defined in [9] in analogy with the classical notion of a degraded broadcast channel [12].

Definition 1 *A channel $\mathcal{N}$ is called* degradable *if there exists a completely positive trace preserving* degrading map $\mathcal{D}$ *such that*

$$\mathcal{D} \circ \mathcal{N} = \widehat{\mathcal{N}},$$

where $\mathcal{N}(\rho) = \mathrm{Tr}_E\, U_{\mathcal{N}}\rho U_{\mathcal{N}}^\dagger$ and $\widehat{\mathcal{N}}(\rho) = \mathrm{Tr}_B\, U_{\mathcal{N}}\rho U_{\mathcal{N}}^\dagger$.

Below, we will prove that the private classical capacity of a degradable channel is equal to its quantum capacity. The quantum capacity is, in turn, equal to the single-letter optimized coherent information. This result is very much in line with the findings of [10], in which it was shown that in the classical case if Z is a degraded version of Y, the noisy processing in Eq. (1) is unecessary.

Theorem 2 *If $\mathcal{N}$ is degradable, then*

$$C_p(\mathcal{N}) = Q^{(1)}(\mathcal{N}) = \sup_\phi I(A\rangle B)_{I\otimes\mathcal{N}|\phi\rangle\langle\phi|},$$

where $I(A\rangle B)_\rho = S(B) - S(AB)$.

To prove this, we will need the following lemma.

Lemma 3 *If $\mathcal{N}$ is degradable, then*

$$C_p^{(1)}(\mathcal{N}) = Q^{(1)}(\mathcal{N}) = \sup_\phi I(A\rangle B)_{I\otimes\mathcal{N}|\phi\rangle\langle\phi|}.$$

Proof Let $\mathcal{N}$ be degradable, fix $\phi = \sum_x p_x|\varphi^x\rangle\langle\varphi^x|$ and let

$$\omega_{XTBE} = \sum_{x,t} p_{x,t}|x\rangle\langle x|_X \otimes |t\rangle\langle t|_T \otimes U_{\mathcal{N}}|\varphi_x\rangle\langle\varphi_x|U_{\mathcal{N}}^\dagger.$$

Then $I(X;B) = I(T;B) + I(X;B|T)$, which is a consequence of the chain rule, together with the fact that $I(XT;B) = I(X;B)$ because $X \to T$. This implies that

$$\begin{aligned} C_p^{(1)}(\mathcal{N}) &= \sup_{\{p_x,|\varphi_x\rangle\}, X\to T} \left(I(T;B)_\omega - I(T;E)_\omega \right) \\ &= \sup_{\{p_x,|\varphi_x\rangle\}, X\to T} \Big(I(X;B) - I(X;E) \\ &\quad - \left(I(X;B|T) - I(X;E|T) \right) \Big). \end{aligned}$$

Since $\mathcal{N}$ is degradable, and conditional mutual information is monotonic under local operations (LO) when the system conditioned on is classical (an immediate consequence of the LO monotonicity of mutual information, itself a consequence of strong subadditivity), we have $I(X;B|T) \geq I(X;E|T)$, so that

$$\begin{aligned} C_p^{(1)}(\mathcal{N}) &= \sup_{\{p_x,|\varphi_x\rangle\}} \left(I(X;B) - I(X;E) \right) \\ &= \sup_{\{p_x,|\varphi_x\rangle\}} \left(S(B) - S(B|X) - S(E) + S(E|X) \right) \\ &= \sup_\phi \left(S(B) - S(E) \right) \\ &= \sup_\phi I(A\rangle B) = Q^{(1)}(\mathcal{N}). \end{aligned}$$

□

The following lemma, which shows that $Q^{(1)}$ is additive for degradable channels, was proved in [9]. We provide an alternate proof for both clarity and completeness.

Lemma 4 *For $\mathcal{N}_1$ and $\mathcal{N}_2$ degradable,*

$$Q^{(1)}(\mathcal{N}_1 \otimes \mathcal{N}_2) = Q^{(1)}(\mathcal{N}_1) + Q^{(1)}(\mathcal{N}_2).$$

Proof Let $|\phi\rangle_{AA_1'A_2'}$ be optimal for $Q^{(1)}(\mathcal{N}_1 \otimes \mathcal{N}_2)$, namely

$$Q^{(1)}(\mathcal{N}_1 \otimes \mathcal{N}_2) = I(A\rangle B_1 B_2)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2|\phi\rangle\langle\phi|}.$$

We would like to show that

$$I(AA_1'\rangle B_2) + I(AA_2'\rangle B_1) \geq I(A\rangle B_1 B_2), \tag{2}$$

since this would immediately imply $Q^{(1)}(\mathcal{N}_1) + Q^{(1)}(\mathcal{N}_2) \geq Q^{(1)}(\mathcal{N}_1 \otimes \mathcal{N}_2)$, and therefore the theorem.

In fact, Eq. (2) is equivalent to

$$I(B_1; B_2) \geq I(E_1; E_2),$$

which, is satisfied due to the degradability of $\mathcal{N}_1$ and $\mathcal{N}_2$ together with the monotonicity of mutual information under local operations. □

We are now in a position to prove Theorem 2.

Proof [of Theorem 2] Let $\mathcal{N}$ be degradable. Then, from [4], the secret-key capacity of $\mathcal{N}$ is

$$C_p(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} C_p^{(1)}(\mathcal{N}^{\otimes n}).$$

By Lemma 3 and the degradability of $\mathcal{N}^{\otimes n}$, we have

$$C_p(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} Q^{(1)}(\mathcal{N}^{\otimes n}),$$

while Lemma 4 gives us $C_p(\mathcal{N}) = Q^{(1)}(\mathcal{N})$. □

III. PRIVATE CLASSICAL CAPACITY WITH A SYMMETRIC SIDE-CHANNEL

Before defining the capacity to be studied, we must first formally define the notion of a private classical code. An (n, K) key code, C, is a set of K states on $A^{\otimes n}$, together with a decoding operation $\mathcal{D}_n : \mathcal{B}(B^{\otimes n}) \to \{1, \ldots K\}$. The rate of such a code is $(\log K)/n$. Such a code is called ϵ-good for a channel $\mathcal{N}^{(n)}$ (mapping $A^{\otimes n}$ to $B^{\otimes n}$) if, defining

$$\rho_{AB^{\otimes n}E^{\otimes n}} = \frac{1}{K}\sum_{x=1}^{K} |x\rangle\langle x|_A \otimes U_{\mathcal{N}^{(n)}}\rho_x(U_{\mathcal{N}^{(n)}})^\dagger,$$

we have

$$||I_A \otimes \mathcal{D}_n \otimes I_{E^{\otimes n}}(\rho_{AB^{\otimes n}E^{\otimes n}}) - \frac{1}{K}\sum_{x=1}^{K} |x\rangle\langle x| \otimes |x\rangle\langle x| \otimes \rho_E||_1 < \epsilon.$$

We say that a rate R is achievable over $\mathcal{N}^{\otimes n}$ if for every $\epsilon > 0$ and all sufficiently large n there is a a code $C_n \subset A^{\otimes n}$ that is ϵ-good for $\mathcal{N}^{\otimes n}$ with $\lim_{n\to\infty} \frac{\log |C_n|}{n} \geq R$. The private classical capacity of $\mathcal{N}$ is then defined as the maximum achievable rate.

Letting $S = S_d \subset \top \otimes \perp$ be the $d(d+1)/2$-dimensional symmetric subspace between $\top$ and $\perp$ and $V_d : \mathbb{C}^{d(d+1)/2} \to S$, we call

$$\mathcal{A}_d(\rho) = \mathrm{Tr}_\perp V_d \rho V_d^\dagger$$

the *d-dimensional symmetric channel*. Note that $\mathcal{A}_d$ maps states on $\mathbb{C}^{d(d+1)/2}$ to states on $\mathbb{C}^d$.

The symmetric side-channel assisted private classical capacity of a channel $\mathcal{N}$ is simply the private capacity of $\mathcal{N}$ when assisted by an arbitrary symmetric channel. More formally, we say that a rate R is ss-achievable if for all $\epsilon > 0$ and sufficiently large n there is a d_n such that R is ϵ-achievable over $\mathcal{N}^{\otimes n} \otimes \mathcal{A}_{d_n}$. The ss-private classical capacity is then the maximum ss-achievable key rate. The main result of this work is the following theorem characterizing the ss-private capacity.

Theorem 5 *The ss-private capacity of $\mathcal{N}$ is*

$$C_{p,ss}^{(1)}(\mathcal{N}) = \sup_{\{p_x, |\varphi_x\rangle_{AFG}\}X\to T} \left(I(T;BF) - I(T;EG)\right), \tag{3}$$

with the optimization over $|\varphi_x\rangle_{AFG}$ symmetric in FG.

Note that this expression for $C_{p,ss}$ is related to but differs from the upper bound presented in [13], which in this case translates to

$$\sup_{\{p_x, |\varphi_x\rangle_A\}X\to\sigma_U, X\to\sigma_V} \left(I(U;BV) - I(U;EV)\right)$$

in that the optimization in Theorem 5 is restricted to classical T rather than a general σ_U. So that besides admitting an operational interpretation, our bound will in general be tighter.

A useful alternative characterization of $C_{p,ss}$ is given by

$$C_{p,ss}^{(1)}(\mathcal{N}) = \sup_d C_p^{(1)}(\mathcal{N} \otimes \mathcal{A}_d), \tag{4}$$

which can be seen to be equivalent to Eq. (3) as follows. To see that Eq. (3) can be no bigger than Eq. (4), note that any ensemble $\{p_x, |\varphi_x\rangle_{AFG}\}$ with $|\varphi_x\rangle_{AFG}$ symmetric in FG can be generated using $\mathcal{A}_d$ with $d = d_F$. Alternatively, given any ensemble of states, $\{p_x, |\varphi_{AS_d}\rangle\}$, where S_d is the input to $\mathcal{A}_d$, we retrieve an ensemble $\{p_x, I \otimes U_{\mathcal{A}_d}|\varphi_{AS_d}\rangle\}$ which is symmetric in FG, so that Eq. (3) is no smaller than Eq. (4).

Before proving the theorem, we provide a multi-letter characterization of the capacity.

Lemma 6

$$C_{p,ss}(\mathcal{N}) = \lim_{n\to\infty} \frac{1}{n} C_{p,ss}^{(1)}(\mathcal{N}^{\otimes n})$$

Proof To see that the ss-private capacity is no less than the right-hand side, note that for any ensemble $\{p_x, |\varphi_x\rangle_{A^nFG}\}$ symmetric in FG and $X \to T$, a rate of

$$\frac{1}{n}\left(I(T;B^nF) - I(T;E^nG)\right)$$

is achievable by the coding theorem of [4].

Conversely, fix $\epsilon > 0$, let $\{\frac{1}{2^{nR}}, \rho^k_{(A')^nFG}\}$ be an (n,ϵ) ss-private code, and let

$$\omega = \frac{1}{2^{nR}}\sum_{k=1}^{2^{nR}} |k\rangle\langle k|_T \otimes \rho^k_{(A')^nFG}.$$

Then, letting $\mathcal{D}$ be the decoding operation associated with the code,

$$\sigma = \left(I_T \otimes U_{\mathcal{N}}^{\otimes n} \otimes I_{FG}\right)\omega\left(I_T \otimes (U_{\mathcal{N}}^{\dagger})^{\otimes n} \otimes I_{FG}\right),$$

and

$$\rho = I_T \otimes \mathcal{D}_{B^n} \otimes I_{EFG}(\sigma),$$

we have

$$\left|\left|\rho - \frac{1}{2^{nR}}\sum_{k=1}^{2^{nR}} |k\rangle\langle k|_T \otimes |k\rangle\langle k|_C \otimes \rho_{EG}\right|\right|_1 \leq \epsilon.$$

As a result,

$$\begin{aligned} I(T;B^nF)_{\sigma_{TB^nF}} &\geq I(T;C)_{\rho_{TC}} \\ &\geq nR - 2\left(2\epsilon nR + H(\epsilon)\right), \end{aligned}$$

where in the last line we have used the continuity result of [14]. Similarly, since $\sigma_{TE^nG} \approx \sigma_T \otimes \sigma_{E^nT}$, we have

$$I(T;E^nG)_{\sigma_{TE^nG}} \leq 2\left(2\epsilon nR + H(\epsilon)\right),$$

so that

$$\begin{aligned} I(T;B^nF)_{\sigma_{TB^nF}} - I(T;E^nG)_{\sigma_{TE^nG}} &\geq nR - 4\left(2\epsilon nR + H(\epsilon)\right) \\ &= nR(1-8\epsilon) - 4H(\epsilon). \end{aligned}$$

Thus,

$$R \leq \frac{1}{1-8\epsilon}\left(\frac{1}{n}C^{(1)}_{p,ss}(\mathcal{N}^{\otimes n}) + 4H(\epsilon)\right).$$

□

Now using the following lemma, which shows that $C^{(1)}_{p,ss}$ is additive, we will be in a position to prove Theorem 5.

Lemma 7 *$C^{(1)}_{p,ss}$ is additive:*

$$C^{(1)}_{p,ss}\left(\mathcal{N}_1 \otimes \mathcal{N}_2\right) = C^{(1)}_{p,ss}\left(\mathcal{N}_1\right) + C^{(1)}_{p,ss}\left(\mathcal{N}_2\right).$$

Proof For any $|\phi_x\rangle_{A_1A_2FG}$ symmetric in FG and $X \to T$, let

$$|\phi^1_x\rangle_{A_1B_2E_2FGC_1C_2} =$$

$$\frac{1}{\sqrt{2}}\Big(I_{A_1} \otimes U_{\mathcal{N}_2} \otimes I_{FG}|\phi_x\rangle|01\rangle_{C_1C_2} +$$

$$\left(I_{A_1} \otimes \mathrm{SWAP}_{B_2E_2} \otimes I_{FG}\right) I_{A_1} \otimes U_{\mathcal{N}_2} \otimes I_{FG}|\phi_x\rangle|10\rangle_{C_1C_2}\Big),,$$

and $\phi^1 = \sum_{x,t} p(x,t)|t\rangle\langle t| \otimes \phi^1_x$.

Then, labeling $\tilde{F}_1 = B_2FC_1$ and $\tilde{G}_1 = E_2GC_2$, we have

$$\begin{aligned} &C^{(1)}_{p,ss}(\mathcal{N}_1) \geq I(T;B_1\tilde{F}_1)_{I\otimes\mathcal{N}_1\otimes I(\phi^1)} - I(T;E_1\tilde{G}_1)_{I\otimes\mathcal{N}_1\otimes I(\phi^1)} \\ &= \frac{1}{2}\Bigg(I(T;B_1B_2F)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} + I(T;B_1E_2F)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} \\ &\quad - I(T;E_1B_2G)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} - I(T;E_1E_2G)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)}\Bigg). \end{aligned}$$

Similarly defining $|\phi^2_x\rangle$, we find

$$\begin{aligned} &C^{(1)}_{p,ss}(\mathcal{N}_2) \geq I(T;B_2\tilde{F}_2)_{I\otimes\mathcal{N}_2\otimes I(\phi^2)} - I(T;E_2\tilde{G}_2)_{I\otimes\mathcal{N}_2\otimes I(\phi^2)} \\ &= \frac{1}{2}\Bigg(I(T;B_1B_2F)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} + I(T;E_1B_2F)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} \\ &\quad - I(T;B_1E_2G)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} - I(T;E_1E_2G)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)}\Bigg), \end{aligned}$$

so that

$$\begin{aligned} &C^{(1)}_{p,ss}(\mathcal{N}_1) + C^{(1)}_{p,ss}(\mathcal{N}_2) \geq \\ &I(T;B_1B_2F)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)} - I(T;E_1E_2G)_{I\otimes\mathcal{N}_1\otimes\mathcal{N}_2\otimes I(\phi)}. \end{aligned}$$

Since this is true for any $|\phi_x\rangle$, we have

$$C^{(1)}_{p,ss}(\mathcal{N}_1) + C^{(1)}_{p,ss}(\mathcal{N}_2) \geq C^{(1)}_{p,ss}(\mathcal{N}_1 \otimes \mathcal{N}_2).$$

□

Proof [of Theorem 5] By Lemma 6, we have

$$C_{p,ss}(\mathcal{N}) = \lim_{n\to\infty}\frac{1}{n}C^{(1)}_{p,ss}(\mathcal{N}^{\otimes n}),$$

whereas Lemma 7 implies $C^{(1)}_{p,ss}(\mathcal{N}^{\otimes n}) = nC^{(1)}_{p,ss}(\mathcal{N})$, which gives the result. □

We now show that $C_{p,ss}$ is convex, a property that the unassisted private classical capacity is not known to possess.

Lemma 8 *$C_{p,ss}$ is convex:*

$$C_{p,ss}((1-p)\mathcal{N}_0+p\mathcal{N}_1) \leq (1-p)C_{p,ss}(\mathcal{N}_0)+pC_{p,ss}(\mathcal{N}_1).$$

Proof Letting $\mathcal{N} = (1-p)\mathcal{N}_0\otimes|0\rangle\langle 0|_{B_2}+p\mathcal{N}_1\otimes|1\rangle\langle 1|_{B_2}$, we consider the purification of $\mathcal{N}$ that gives Eve the which channel information in a system E_2. Noting that for any $\mathcal{N}$ and $\mathcal{M}$, $C_{p,ss}(\mathcal{N}) \geq C_{p,ss}(\mathcal{M}\circ\mathcal{N})$, we have

$$\begin{aligned}
& C_{p,ss}\left((1-p)\mathcal{N}_0 + p\mathcal{N}_1\right) \\
\leq\ & C_{p,ss}\left((1-p)\mathcal{N}_0\otimes|0\rangle\langle 0|_{B_2} + p\mathcal{N}_1\otimes|1\rangle\langle 1|_{B_2}\right) \\
=\ & \sup_{\{p_x,|\varphi^x\rangle\},X\to T}\left(I(T;BB_2F)-I(T;EE_2G)\right) \\
=\ & \sup_{\{p_x,|\varphi^x\rangle\},X\to T}\left(\sum_{\alpha=0}^{1} p_\alpha\left[S(T|EG,\alpha)-S(T|BF,\alpha)\right]\right) \\
\leq\ & \sum_{\alpha=0}^{1} p_\alpha \sup_{\{p_x,|\varphi^x\rangle\},X\to T}\left(S(T|EG,\alpha)-S(T|BF,\alpha)\right) \\
=\ & \sum_{\alpha=0}^{1} p_\alpha C_{p,ss}(\mathcal{N}_\alpha),
\end{aligned}$$

where the $|\varphi^x\rangle$ are taken to be on $A'FG$ and symmetric in FG throughout. □

Finally, we demonstrate that the ss-private-capacity of a degradable channel is, in fact, the single-letter optimized coherent information.

Lemma 9 *If $\mathcal{N}$ is degradable, $C_{p,ss}(\mathcal{N}) = Q^{(1)}(\mathcal{N})$.*

Proof For any d and degradable $\mathcal{N}$, it is also the case that $\mathcal{N}\otimes\mathcal{A}_d$ is degradable. As a result,

$$\begin{aligned}
C_p^{(1)}(\mathcal{N}\otimes\mathcal{A}_d) &= Q^{(1)}(\mathcal{N}\otimes\mathcal{A}_d) \\
&= Q^{(1)}(\mathcal{N}) + Q^{(1)}(\mathcal{A}_d) = Q^{(1)}(\mathcal{N}),
\end{aligned}$$

so that, by the characterization of $C_{p,ss}$ in Eq.(4), we have $C_{p,ss}(\mathcal{N}) = Q^{(1)}(\mathcal{N})$. □

We will now use the convexity of $C_{p,ss}$ to show that the following quantity, which we call the *cost of degradable mixing*, is an upper bound for C_p.

Definition 10 *We define the* cost of degradable mixing *as*

$$C_{\mathrm{DM}}(\mathcal{N}) = \inf_{\{p_i,\mathcal{N}_i,\mathcal{D}_i\}}\sum_i p_i Q^{(1)}(\mathcal{N}_i),$$

where the infimum is over $\{p_i,\mathcal{N}_i,\mathcal{D}_i\}$ such that

$$\mathcal{N} = \sum_i p_i\mathcal{D}_i\circ\mathcal{N}_i$$

and each $\mathcal{N}_i$ is either degradable or anti-degradable.

That is, we will prove the following theorem.

Theorem 11 *The cost of degradable mixing of a quantum channel is an upper bound for its private classical capacity. In other words, $C_p(\mathcal{N}) \leq C_{\mathrm{DM}}(\mathcal{N})$.*

Notice that, by restricting our $\mathcal{N}_i$ to be either the identity channel or be both degradable and anti-degradable, we would recover the upper bound of [15].

Proof Let $\mathcal{N} = \sum_i p_i\mathcal{D}_i\circ\mathcal{N}_i$ be a decomposition of $\mathcal{N}$ with each $\mathcal{N}_i$ either degradable or anti-degradable. Then, noting that $C_p(\mathcal{N}) \leq C_{p,ss}(\mathcal{N})$, and using the convexity of $C_{p,ss}$, we have

$$\begin{aligned}
C_p(\mathcal{N}) &\leq C_{p,ss}\left(\sum_i p_i\mathcal{D}_i\circ\mathcal{N}_i\right) \\
&\leq \sum_i p_i C_{p,ss}\left(\mathcal{D}_i\circ\mathcal{N}_i\right) \\
&\leq \sum_i p_i C_{p,ss}\left(\mathcal{N}_i\right) \\
&= \sum_i p_i Q^{(1)}(\mathcal{N}_i),
\end{aligned}$$

where in the last line we have used the fact that for $\mathcal{N}_i$ either degradable or antidegradable, $C_{p,ss}(\mathcal{N}_i) = Q^{(1)}(\mathcal{N}_i)$. □

One might wonder about the inclusion of $\mathcal{D}_i$s in the definition of the cost of degradable mixing—wouldn't the bound be tighter if they were all chosen to be the identity? The trouble is that not all channels can be written as a convex combination of degradable and anti-degradable channels, but *any* channel can be decomposed into the form required by our definition (e.g., choose only one term, and let $\mathcal{N}_1 = I$ and $\mathcal{D}_1 = \mathcal{N}$, the channel of interest). In particular, while all extremal qubit channels are either degradable or antidegradable (or both)[20], and therefore any qubit channel can be written as a convex combination of such channels, the same is not true in higher dimension. For example[27], the tensor product of two extremal qubit channels, one degradable and the other anti-degradable (but neither both), is generically an extremal channel on two qubits, but is neither degradable nor anti-degradable, and in light of its extremality cannot be decomposed into such channels. To get around this, we include the $\mathcal{D}_i$s in the definition of the cost of degradable mixing. In the two qubit example, we find C_{DM} is exactly equal to the quantum capacity of the degradable channel, and therefore so is the private capacity, incidentally providing an example of a nondegradable channel for which the private and quantum capacities coincide.

IV. SOME SPECIFIC CHANNELS

Theorem 11 gives us a technique for bounding the private capacity of a general channel, $\mathcal{N}$, in terms of the private

capacity of a set of degradable channels appearing in a convex decomposition of $\mathcal{N}$. We now use this method to provide upper bounds for the key capacity of two channels of interest for quantum key distribution — the Pauli channel with independent phase and amplitude noise and the depolarizing channel. The resulting bounds meet or exceed all previously known bounds on the private classical capacity of these channels [15, 16, 17, 18].

A. Degradable Channels

In this subsection we explicitly evaluate the private capacity of some degradable channels.

A qubit channels with two Kraus operators has, up to local unitaries, Kraus operators equal to [19]

$$A_0 = \begin{pmatrix} \sqrt{1-\gamma} & 0 \\ 0 & \sqrt{1-\delta} \end{pmatrix} \quad A_1 = \begin{pmatrix} 0 & \sqrt{\delta} \\ \sqrt{\gamma} & 0 \end{pmatrix}.$$

It was shown in [20] that any such channel is either degradable or anti-degradable. The private capacity of such a channel is thus the optimized single-letter coherent information:

$$C_p(\mathcal{N}_{(\gamma,\delta)}) = \max_{t\in[0,1]} \left[H(t(1-\gamma)+(1-t)\delta) - H(t\gamma+(1-t)\delta)\right].$$

This result includes the dephasing and amplitude damping channels as a special case: setting $\gamma = 0$ gives an amplitude damping channel, whereas setting $\gamma = \delta$ gives the bitflip channel (which is unitarily equivalent to a dephasing channel)[28].

The erasure channel with erasure probability p, which maps $\mathbb{C}^d$ into $\mathbb{C}^d \oplus |e\rangle$, acts as

$$\mathcal{N}^{\text{erasure}}_{(p,d)}(\rho) = (1-p)\rho + p|e\rangle\langle e|.$$

This channel is also degradable, and as a result its private classical capacity is exactly equal to its quantum capacity:

$$C_p(\mathcal{N}^{\text{erasure}}_{(p,d)}) = (1-2p)\log d.$$

B. Independent Phase and Amplitude errors

The Pauli channel with independent amplitude and phase noise is an interesting case because of its relation to BB84, and also because it's easy to write as a convex combination of degradables—it's just an equal mixture of two amplitude damping-type channels.

Written explicitly, the channel we are considering is

$$\begin{aligned}\mathcal{N}_{(q(1-q),q^2,q(1-q))}(\rho) \;=\; & (1-q(2-q))\rho + q(1-q)X\rho X \\ & + q^2 Y\rho Y + q(1-q)Z\rho Z,\end{aligned}$$

which can also be written as

$$\frac{1}{2}U\mathcal{N}^{\text{ampdamp}}_{\gamma_q}(U^\dagger\rho U)U^\dagger + \frac{1}{2}UX\mathcal{N}^{\text{ampdamp}}_{\gamma_q}(XU^\dagger\rho UX)XU^\dagger,$$

where $U = e^{i\frac{\pi}{4}X}$ and $\gamma_q = 4q(1-q)$.

From the previous subsection, the private capacity of an amplitude damping channel with noise parameter γ is

$$f(\gamma) = \max_{t\in[0,1]} \left(H(t(1-\gamma)) - H(t\gamma)\right),$$

so that

$$C_p(\mathcal{N}_{(q(1-q),q^2,q(1-q))}) \leq f(\gamma_q).$$

This gives a threshold of $\frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)$ beyond which no key can be generated, which is the same as found for BB84 in[16] , and also confirmed in [13] and [15]. We can also write the $\mathcal{N}_{(q(1-q),q^2,q(1-q))}$ as a convex combination of dephasing channels with dephasing probability $q(2-q)$, which results in slightly tighter bounds for very small noise (i.e., $q < 0.02$). Our combined upper bound on key rate is given by

$$C_p(\mathcal{N}_{(q(1-q),q^2,q(1-q))}) \leq \text{conv}\left(1 - H(q(2-q)), f(\gamma_q)\right), \tag{5}$$

and tightens the previous best bounds of [16] (which considered only protocols without noisy processing), and the (straight line) bound found in [15] for all $0 < q < \frac{1}{2}\left(1-\frac{1}{\sqrt{2}}\right)$ (see Figure 1). The fact that we surpass the bound of [16] is particularly interesting, since it is also an *achievable* key rate against an adversary restricted to individual attacks. Our bound thus shows that a completely general attack is strictly stronger than an individual attack.

C. Depolarizing channel

A depolarizing channel with error probability p is a convex combination of six amplitude damping channels with error parameter

$$\eta_p = 4\sqrt{1-p}\left(1-\sqrt{1-p}\right)$$

(see [5] for details). It is also a convex combination of three dephasing channels with error probability p. Finally, the secret key capacity is zero whenever a channel is antidegradable, which happens at $p = 1/4$ [21], so that the convexity of $C_{p,ss}$ then implies

$$C_{p,ss}(\mathcal{N}_p) \leq \text{conv}\left(1 - H(p), f(\eta_p), (1-4p)_+\right),$$

where we have let $x_+ = x$ if $x > 0$ and 0 otherwise. This expression is equal to the upper bound on the depolarizing channel's *quantum* capacity found in [5], so that

the best known upper bounds for this channel actually coincide.

It is worth mentioning that the bound on the threshold for the six-state protocol reported in [13] is strictly stronger than the $p = 1/4$ threshold implied by our bound. However, the [13] bound *does not* apply to the private capacity of the depolarizing channel, since it is valid only for a restricted set of input states.

For comparison with the QKD literature, note that the relationship between quantum bit error rate, q, and depolarizing error probability, p, is $q = 2p/3$.

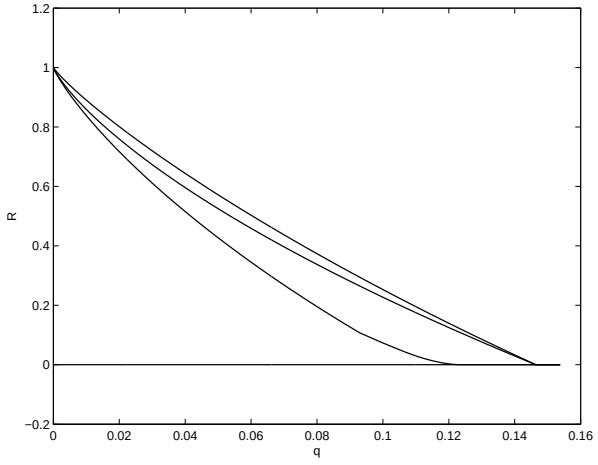

FIG. 1: Bounds on the key rate of BB84 with one-way post-processing as a function of quantum bit error rate, q. The lower curve is the best known achievable key rate from [11, 13]. The upper curve is the "optimal eavesdropping"' bound on BB84 (*without* noisy processing) found in [16], representing the best possible individual attack. The middle curve is our upper bound from Eq. (5).

D. Pauli Channel

For a general Pauli channel we find the following bound.

Theorem 12 *Let*

$$\mathcal{N}_{\mathbf{p}}(\rho) = (1 - |\mathbf{p}|)\rho + p_1 X\rho X + p_2 Y\rho Y + p_3 Z\rho Z.$$

Then the private classical capacity of $\mathcal{N}_{\mathbf{p}}$ *satisfies*

$$C_p(\mathcal{N}_{\mathbf{p}}) \leq 1 - H(|\mathbf{p}|),$$

where $|\mathbf{p}| = p_1 + p_2 + p_3$.

Proof Letting $\alpha_i = p_i/|\mathbf{p}|$, we have

$$\mathcal{N}_{\mathbf{p}}(\rho) = \alpha_1 \mathcal{N}^X_{|\mathbf{p}|}(\rho) + \alpha_2 \mathcal{N}^Y_{|\mathbf{p}|}(\rho) + \alpha_3 \mathcal{N}^Z_{|\mathbf{p}|},$$

where we have let $\mathcal{N}^X_p(\rho) = (1-p)\rho + pX\rho X$, and similarly for $\mathcal{N}^Y_p$ and $\mathcal{N}^Z_p$. This is a convex combination of dephasing-like channels with error probability $|\mathbf{p}|$, which are degradable and have a private capacity of $1 - H(|\mathbf{p}|)$, so that by Theorem 11 we have the result. □

It is not entirely clear how to best decompose a Pauli channel into a convex combination of amplitude damping channels, but it seems likely that such a decomposition (or perhaps a decomposition into channels with two Kraus operators) would outperform the current bound for high noise levels.

E. Relationship to collective attacks in QKD

In this subsection we describe how the above upper bounds on the private capacity correspond to collective attacks on quantum key distribution protocols. Consider the decomposition of a channel $\mathcal{N}$ into a convex combination of degradable channels, $\mathcal{N}_i$:

$$\mathcal{N}(\rho) = \sum_i p_i \mathcal{N}_i,$$

where we will call the isometric extension of $\mathcal{N}_i$ $U_{\mathcal{N}_i} : A \to BE$. The attack associated with this decomposition is as follows: For each signal state sent, Eve applies $U_{\mathcal{N}_i}$ with probability p_i, sends the B system to Bob, and stores her various E systems until the end of the protocol. After the protocol is complete, Eve collects all of her E systems associated with $\mathcal{N}_i$ and (jointly) measures which of the typical eigenvectors of $\rho_{E_i}^{\otimes p_i n}$ the state is in. Because $\mathcal{N}_i$ is degradable, we can calculate exactly how much secret key Alice and Bob can distill from the resulting state—they can get a key rate of exactly $Q^{(1)}(\mathcal{N}_i)$. Because a fraction p_i of the signal states are subjected to $\mathcal{N}_i$, the overall key rate is no more than $\sum_i p_i Q^{(1)}(\mathcal{N}_i)$.

V. DISCUSSION

We have studied the capacity of a quantum channel for private classical communication when assisted by symmetric channel of an arbitrary size. For a general channel, we found a single letter formula that, unfortunately, involves an optimization over an auxiliary space that is a prior unbounded. For degradable channels, we further showed that this optimization can be performed explicitly, and in fact the ss-private capacity of such a channel is exactly equal to its single-letter optimized coherent information. Using this fact, together with the convexity of the ss-capacity for general channels, we showed how to find upper bounds on the (unassisted) private capacity of a general channel, and provided such bounds for two channels of interest for quantum key distribution.

The most important question we have left unanswered is whether it is possible to bound the dimension of the

symmetric channel necessary to achieve the optimum of the ss-capacity formula found in Theorem 5. This could allow very tight bounds on the unassisted capacity. In fact, we are unaware of any channel for which the ss-private capacity and unassisted private capacity differ, and the conjecture that they are the same is equivalent to the additivity of the unassisted capacity, C_p.

We note that for both the independent amplitude and phase noise and the depolarizing channel, the upper bounds are the convex hull of a bound based on decomposition into dephasing channels, which is strongest in the low noise regime, and a decompostion into amplitude damping channels, which is stronger in the high noise regime. This suggests that the best collective attacks on quantum key distribution protocols will be qualitatively different in the high and low noise regimes.

It is an interesting question whether there are zero capacity degradable channels that are not antidegradable. This possibility is intriguing, since the best known bounds on the zeros of the capacity of most channels come from a no-cloning argument (which is essentially to observe that the channel is antidegradable), but these bounds are usually not particularly close to the corresponding lower bounds. Such a channel would also be useful for improving estimates on C_{DM} for general channels.

Finally, this work demonstrates (along with [5]) that assistance from a symmetric side channel is a "nice" resource, in the sense that it provides a marked simplification over the unassisted case for the private capacity. Further examples of nice resources are free EPR pairs, which lead to the single-letter formula for the entanglement assisted capacity of [22, 23], and PPT-preserving operations, which dramatically simplify the theory of entanglement manipulations [24, 25]. What are the other "nice" resources?

Acknowledgments

I am grateful to Debbie Leung, John Smolin, Andreas Winter, and Charlie Bennett for helpful conversations, to the Institute for Quantum Computing at the University of Waterloo, where this work was initiated, and the United Kingdom Engineering and Physical Sciences Research Council for financial support.

[1] C. H. Bennett and G. Brassard, Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing p. 175 (1984).
[2] R. Renner (2005), Ph.d. Thesis, Swiss Federal Institute of Technology.
[3] D. Gottesman and H.-K. Lo, IEEE Trans. Inf. Theory **49**, 457 (2003).
[4] I. Devetak, IEEE Trans. Inf. Theory **51**, 44 (2005), arXiv:quant-ph/0304127.
[5] G. Smith, J. Smolin, and A. Winter, arXiv:quant-ph/0607039.
[6] I. Devetak and A. Winter, Phys. Rev. Lett. **93**, 080501 (2004), arXiv:quant-ph/0307053.
[7] B. Schumacher and M. D. Westmoreland, Phys. Rev. Lett. **80**, 5695 (1998).
[8] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim, Phys. Rev. Lett. **94**, 160502 (2005), arXiv:quant-ph/0309110.
[9] I. Devetak and P. W. Shor, Comm. Math. Phys. **256**, 287 (2005), arXiv:quant-ph/0311131.
[10] I. Csiszar and J. Korner, IEEE Trans. Inf. Theory **24**, 339 (1978).
[11] G. Smith, J. Renes, and J. A. Smolin, arXiv:quant-ph/0607018.
[12] T. Cover, IEEE Trans. Inf. Theory **18**, 2 (1972).
[13] B. Kraus, N. Gisin, and R. Renner, Phys. Rev. Lett. **95**, 080501 (2005), arXiv:quant-ph/0410215.
[14] R. Alicki and M. Fannes, arXiv:quant-ph/0312081.
[15] T. Moroder, M. Curty, and N. Lutkenhaus, arXiv:quant-ph/0603270.
[16] C. A. Fuchs, N. Gisin, R. B. Griffiths, C.-S. Niu, and A. Peres, Phys. Rev. A **56**, 1163 (1997).
[17] D. Bruss, Phys. Rev. Lett. **81**, 3018 (1998).
[18] H. Bechmann-Pasquinucci and N. Gisin, Phys. Rev. A **59**, 4238 (1999).
[19] M. B. Ruskai, S. Szarek, and E. Werner, Lin. Alg. .Appl. **347**, 159 (2002), arXiv:quant-ph/0101003.
[20] M. Wolf and D. Perez-Garcia, Phys. Rev. A **75**, 012303 (2007), arXiv:quant-ph/0607070.
[21] D. Bruss, D. P. DiVincenzo, A. Ekert, C. A. Fuchs, C. Macchiavello, and J. A. Smolin, Phys. Rev. A **57**, 2368 (1998), arXiv:quant-ph/9705038.
[22] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, IEEE Trans. Inf. Theory **48**, 2637 (2002).
[23] C. Adami and N. J. Cerf, Physical Review A **56**, 3470 (1997), arXiv:quant-ph/9609024.
[24] T. Eggeling, K. Vollbrecht, R. F. Werner, and M. M. Wolf, Phys. Rev. Lett. **87**, 257902 (2001), arXiv:quant-ph/0104095.
[25] E. M. Rains, IEEE Trans. Inf. Theory **47**, 2921 (2001).
[26] R. Renner and R. Koenig, Proc. of TCC, LNCS, Springer **3378** (2005), arXiv: quant-ph/0403133.
[27] Thanks to Debbie Leung for providing this example.
[28] Note that in [26] it was shown that the optimal key rate achievable for dephasing noise of rate p on a maximally correlated classical string is $1 - H(p)$, but because they do not consider general signal states, this does not quite show that the private classical capacity of the dephasing channel is the same value, though this formula is implied by our result.